ORIGINAL ARTICLE

Networks

The Complexity Landscape of Disaster-Aware Network Extension Problems

Balázs Vass[1,2] | Beáta Éva Nagy[1] | Balázs Brányi[1] | János Tapolcai[1]

[1] High-Speed Networks Laboratory (HSN*Lab*), Department of Telecommunication and Media Informatics, Faculty of Electrical Engineering and Informatics (VIK), Budapest University of Technology and Economics (BME), MTA-BME Future Internet Research Group, and ELKH-BME Information Systems Research Group, 1111, Budapest, Hungary

[2] Faculty of Mathematics and Computer Science, Babeș-Bolyai University of Cluj Napoca, Romania

Correspondence
Balázs Vass
Email: vb@tmit.bme.hu

Funding information
This work was partly supported by Project no. 135606, Project no. 134604, Project no. 128062, and Project no. 137698, which have been implemented with support from the National Research, Development and Innovation Fund of Hungary, financed under the ANN_20, FK_20, K_18, and PD_21 funding schemes, respectively. Supported by the ÚNKP-23-4-II-BME-345 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.

This paper deals with the complexity of problems related to finding cost-efficient, disaster-aware cable routes. We overview various mathematical problems studied to augment a backbone network topology to make it more robust against regional failures. These problems either consider adding a single cable, multiple cables, or even nodes too. They adapt simplistic or more sophisticated regional failure models. Their objective is to identify the network's weak points or minimize the investment cost concerning the risk of a network outage. We investigate the tradeoffs in mathematical modeling for the same real-world scenario, where more sophisticated models face more computationally challenging problems. We have seen how efficiently computational geometry algorithms can be used to solve simplified problems even for sufficiently large networks. In this paper, we aim to understand why different mathematical models formulated for the same real-world scenario can or cannot be solved efficiently. In particular, we show simplistic mathematical models that formulate NP-hard problems.

KEYWORDS
network augmentation, large scale disasters, regional failure, network robustness, complexity theory, geometric graphs

1 | INTRODUCTION

Several studies revealed how vulnerable the Internet backbone is to natural disasters, such as earthquakes, hurricanes, and tsunamis. Such disasters may destroy several nodes and links located in possibly a few hundred kilometers wide geographic areas [20, 7, 9, 11, 4, 10, 16, 19, 24, 13, 8, 23, 29]. Consequently, extensive research has been conducted to address the challenge of deploying new network equipment, specifically fiber cables [18, 31, 26], with the aim of enhancing network resilience against natural disasters. Broadly speaking, these studies aimed to determine the optimal placement of new resources to maximize the network's ability to survive natural disasters. While these studies had the same final goal, researchers had flexibility in formulating the mathematical models for their respective approaches. Although the differences may seem technical, they have a significant impact on the computational complexity of the underlying algorithms. As a result, some problems can be efficiently solved, while others are intractable.

Inspired by these observations, this paper is dedicated to exploring the factors that contribute to the difficulty of solving such problems in this context. We begin by identifying three primary criteria that are adopted by the mathematical models in this domain:

1. **Network Extension Strategy**: Some models focus on adding a **single cable** to maximize network robustness. Others consider the addition of **multiple cables** in a single step. Alternatively, they go a step further by adding **multiple cables and nodes** simultaneously.
2. **Failure Correlation**: Failures can be treated as **non-correlated**, where the risk of cable cut is solely dependent on its location, and the failures are assumed to be independent of each other. Another approach is considering them **correlated**, where a more complex regional failure model is employed to capture interdependencies among failures.
3. **Main Cost Factor**: Some models adopt a **worst-case scenario** approach, aiming to identify the network's vulnerabilities and add new links to mitigate them. Others take a more holistic perspective by considering the **probability of disasters**, seeking a cost-efficient solution that balances investment costs and the risk of network outages.

There are other differences between models, such as considering links to be directed or bidirectional, straight lines or polygonal chains. However, our primary focus lies on the categories mentioned above, considering the other differences as consequential.

The computational complexity of the problems in this domain varies significantly. Some problems can be solved in polynomial time (i.e., they are in $\mathcal{P}$) using computational geometry tools, while others are known to be $\mathcal{NP}$-hard. It is important to note that solving $\mathcal{NP}$-hard problems on classical computers is not possible in polynomial time unless $\mathcal{P}=\mathcal{NP}$.

In prior works, researchers have made significant contributions by addressing the complexity of several problems, which are summarized in Table 1. In this paper, our primary objective is to determine the boundary between problems that can be solved in polynomial time and those that are $\mathcal{NP}$-hard. We explore the trade-off between over-simplification and intractability, in the following way. Overly simplistic models may overlook important real-world considerations, while more detailed models introduce computationally demanding problems. In essence, although simplistic problems often fall within the complexity class $\mathcal{P}$, their optimal solutions may not be practically viable. On the other hand, detailed problems are typically $\mathcal{NP}$-hard and require careful input data handling. For instance, cost-efficient models often require extensive empirical hazard data and employ complex algorithms due to their $\mathcal{NP}$-hard nature.

An earlier version of the paper appeared at IEEE RNDM 2022 [28].

TABLE 1 Main contribution of papers considering worst-case models for planar graphs.

	disaster type	existing edges	added edges	adding new nodes	main result on making the network disaster resilient the cheapest way
Akitaya et al. [2]	point	line segments	line segments	not allowed	connected graphs: polynomial algorithm disconnected graphs: $\mathcal{NP}$-hardness
Tapolcai et al. [25]	circular disk	polygonal chains	chains of line segments and circular arcs	not allowed	heuristic algorithms
Thió et al. [3]	horizontal line segment	polygonal chains	polygonal chains	not allowed	heuristic algorithms
The current paper	point or circular disk or horiz. line segment	(no existing edges)	chains of line segments and circular arcs	either allowed or not allowed	disconnected graphs: $\mathcal{NP}$-hardness

We would like to highlight that this paper is an extended and thoroughly revised version of our paper [28]. In the original study, we put more emphasis on comparing the results of two recent papers [25, 22] and investigating the complexities of the problems tackled by them. The current paper has a larger scope, summarizing and charting the complexity results of related prior works (see Tables 1 and 2). In particular, we give a more detailed description of worst-case models. Regarding disaster region shapes, on top of point and circular disk failures, in this study, we also tackle horizontal line segment failures. Additionally, we extend the proofs of our main results (listed in Sections 3.2 and 3.3), by giving better upper bounds in our calculations.

The paper is organized as follows. Section 2 defines the model and problems for the worst-case scenario. Section 3 is the main contribution, proving the $\mathcal{NP}$-hardness of the previously defined problems. Next, Section 4 discusses the cost-efficient models. Section 5 discusses how the rasterization of the map affects the quality of the solution. Finally, Section 6 concludes the paper.

2 | WORST-CASE MODELS AND RELATED PROBLEMS

Throughout Sections 2 and 3, we focus on worst-case analysis. In papers considering worst-case problems [2, 25, 3, 28], the network is modeled as an undirected connected geometric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $n = |\mathcal{V}|$ nodes and $m = |\mathcal{E}|$ edges. The nodes are embedded as points in the Euclidean plane. Each edge is considered a finite sequence of line segments. We suppose there is no link crossing, i.e., no two links share common points except their endpoints (at the nodes). In some problems, each edge of $\mathcal{E}$ is straight. In these cases, we call $\mathcal{G}$ a Planar Straight Line Graph [2].

In real life, disasters can come in various shapes. Most of the problems in Sections 2 and 3 utilize a specific failure model, the *disk failure model with fixed radius* [25]. Here, a regional failure occurs at a point known as the epicenter, which corresponds to an area in the shape of a circular disk c of radius r. Some studies, like [2], simply assume that failure c is a single point on the plane. This would vaguely correspond to a circular disk failure of size $r \searrow 0$. More theoretical works, like [3], deal with horizontal segment failures [1], where the shape of the failure is a horizontal line segment of length r.

Definition 1 *A circular disk* ***failure*** c ***hits*** *an edge e if e intersects the interior of disk c. Similarly, node v is hit by failure c if it is in the interior of c. Let $\mathcal{E}_c$ (and $\mathcal{V}_c$) denote the set of edges (and nodes) hit by disk c.*

[1] We believe the two models are closely related, because a circular disk, after scaling down along its vertical axis, almost looks like a horizontal line segment. Intuitively, the problem is invariant to a transformation of the topology and the shape of the disaster in the same way. As a result, dealing with circular disk failures in a topology that has been stretched vertically is equivalent to handling failures in the shape of horizontal line segments in the original (unscaled) topologies.

All network elements intersecting with the disaster are destroyed, and all other network elements remain unaffected. The goal is to ensure that the network survives the failure.

Definition 2 *A* ***network survives*** *a circular disk failure c if the graph* $\mathcal{G}_c = (\mathcal{V} \setminus \mathcal{V}_c, \mathcal{E} \setminus \mathcal{E}_c)$ *is connected.*

The real-world probabilistic nature of disasters is ignored, the solution should survive all possible disasters. Therefore, the goal is to create a topology that is more likely to remain connected after a natural disaster.

The task is to extend the network topology with one or multiple new links, and in some problems, with new nodes as well. The output of each extension algorithm is a set of new links $\mathcal{E}^+$, with the exact location of their endpoints and the routes of the cables (between the endpoints). If we also allow adding a set of new nodes $\mathcal{V}^+$, their locations also form part of the output. The objective is to minimize the installation cost of the new cables (and nodes if applicable). This cost is related to the total cable length $c(\mathcal{E}^+)$, plus the installation cost $c(\mathcal{V}^+)$ of new nodes if applicable. The cost of implementing a new cable is the same everywhere in the network, i.e., it is directly proportional to its total physical length.

Note that we also consider the case of disconnected graphs, when we build a more resistant topology from scratch. In these cases, all nodes and edges count as new components, and we want to minimize the overall installation costs.

For many disasters, the damage is mainly determined by the physical distance to the epicenter, e.g., attacks via mass destruction weapons. However, some natural disasters cause a damaged area of a specific shape; e.g., tsunamis damage areas near the seashore, earthquakes are likely to happen in seismic zones, and flooding can happen along a riverside [14]. Nevertheless, when preparing the network for the worst case, it is reasonable to assume that nodes/links may be affected by the failure if they are physically close to the disaster's epicenter. It is a reasonable requirement to ensure connectivity for nodes sufficiently distant from the center of the disaster. In other words, maintaining service continuity on the border of the catastrophic region is outside the scope of these studies.

A common assumption of these studies is that they do not consider the limitation of routing. Instead, they focus on whether the topology graph remains connected after the failure. That is, if the network remains connected after the failure, it is assumed that the nodes will be able to communicate.

2.1 | Mathematical Problem Definitions

In the following, multiple problems are defined. The main differentiation is made on whether we allow adding a single cable between two given nodes of the graph, multiple cables between different nodes, or even new nodes as well. In addition to this, the disaster might be a single point of the plane, or a connected set of points (called a "connected regional failure"). The latter incorporates circular disk failures and horizontal line segment failures. The original graph $\mathcal{G}$ is either connected or disconnected. It will turn out that some proofs of $\mathcal{NP}$-hardness require supposing $\mathcal{G}$ being disconnected.

Although the complexities of several problems have already been proved, there are still many uninvestigated ones. In the scope of this paper, we define and prove the complexities of four problems, each based on disconnected input graphs. The aim is to find the minimum cost network topology addition in order to survive a single point or circular disk failure of radius r at any location. These connections are reached by either allowing only new edges, or extra nodes as well. Such extensions result in increased complexity. We will prove in the following, that the decision versions of the problems are $\mathcal{NP}$-hard, and no Fully Polynomial Time Approximation Scheme exists for them (unless $\mathcal{P}$=$\mathcal{NP}$).

Next, we define the problems. For the Network Augmentation problems (where only new edges can be added), the input is the same network topology:

Input: A network represented by an undirected geometric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $n = |\mathcal{V}|$ nodes, that are embedded as points, and the edges as line segments in the Euclidean plane, a maximum radius r of a possible regional failure, and the limit on the total length of the new curves B.

The problem outputs are given as follows:

Network Augmentation (NA) Problem

Output: Are there curves added as new edges $\mathcal{E}^+$ to $\mathcal{E}$, such that the network survives any point failure, and the total length of the curves $c(\mathcal{E}^+)$ is at most B?

Geometric Network Augmentation (GNA) Problem

Output: Are there curves added as new edges $\mathcal{E}^+$ to $\mathcal{E}$, such that the network survives any circular disk failure of radius r (see Definition 2), and the total length of the curves $c(\mathcal{E}^+)$ is at most B?

For the Network Extension problems (where new edges and nodes can be added as well), the input is the same network topology. The cost function is slightly different since we have to factor in the cost of installing new nodes too:

Input: A network represented by an undirected geometric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $n = |\mathcal{V}|$ nodes, that are embedded as points, and the edges as line segments in the Euclidean plane, a maximum radius r of a possible regional failure, and the total cost budget B on installing new network equipment, that is the sum of the total length of the curves $c(\mathcal{E}^+)$ and the installation cost of new nodes $c(\mathcal{V}^+) \geq 0$.

The problem outputs are given as follows:

Network Extension (NE) Problem

Output: Are there nodes $\mathcal{V}^+$ added to $\mathcal{V}$ (so-called Steiner nodes) and curves through $\mathcal{V} \cup \mathcal{V}^+$ added as new edges $\mathcal{E}^+$ to $\mathcal{E}$, such that the network survives any point failure, and the total cost $c(\mathcal{E}^+) + c(\mathcal{V}^+) \leq B$?

Geometric Network Extension (GNE) Problem

Output: Are there nodes $\mathcal{V}^+$ added to $\mathcal{V}$ and curves through $\mathcal{V} \cup \mathcal{V}^+$ added as new edges $\mathcal{E}^+$ to $\mathcal{E}$, such that the network survives any circular disk failure of radius r, and the total cost $c(\mathcal{E}^+) + c(\mathcal{V}^+) \leq B$?

Table 2 summarizes the differences in the problems investigated in the related papers regarding worst-case models and the current study, along with their computational complexities. In the next section, we discuss these results in detail.

3 | COMPLEXITY OF PROBLEMS BASED ON WORST-CASE MODELS

In this section, we cover up the complexity landscape of problems based on worst-case models. In Section 3.1, we present some positive results on the very constrained settings where only one or two links can be added to the existing network. Sections 3.2-3.4 present and discuss our main results, which are the $\mathcal{NP}$-hardness proofs of the problems defined in Section 2.1 (as listed in the last two rows of Table 2). These proofs of $\mathcal{NP}$-hardness are inspired by theoretical papers [12, 17]. Very intuitively, all of our proofs are based on the fact that finding a Hamiltonian

TABLE 2 Complexity of different problems for uniform cable cost. $\mathcal{G}$ is assumed to be a planar graph with links embedded as polygonal chains, and new links being chains of line segments and circular arcs.

Extend with \ Graph & failure	point failure		connected regional failure	
	$\mathcal{G}$ connected	$\mathcal{G}$ disconnected	$\mathcal{G}$ connected	$\mathcal{G}$ disconnected
a single cable	trivial	no solution exists	$\in \mathcal{P}$ Proposition 1	no solution exists
multiple cables	$\in \mathcal{P}$ by Proposition 2 for line segment edges	Problem NA is $\mathcal{NP}$-hard by Lemma 5	open	Problem GNA is $\mathcal{NP}$-hard by Theorem 6
multiple cables & nodes	open	Problem NE is $\mathcal{NP}$-hard by Lemma 7	open	Problem GNE is $\mathcal{NP}$-hard by Lemma 8

circuit in a *grid graph* (Definition 3) is an $\mathcal{NP}$-hard problem. If we take the nodes (without edges) of a grid graph accompanied with a small enough disaster radius, deciding whether a best solution for deploying cables is cheaper than a given threshold translates to answering the question of whether there exists a Hamiltonian circuit in the original grid graph. For an introduction to complexity theory and approximation algorithms, we refer the reader to [6] and [30], respectively. Finally, in Section 3.5, we present our thoughts on the computational complexity of worst-case problems where disasters are constrained to be horizontal line segments.

3.1 | Adding One or Two Links in the Worst-Case Models

The problems in the first row of Table 2 (considering adding a single cable) can be reduced to connecting two given nodes of the network after a point failure or connected regional failure. [25, rephrased Claim 4.] proves that if the input graph is connected, it is possible to find the optimal solution in polynomial time, given that a single cable is added (see the upcoming Proposition 1). Note that if G is not connected, we need to add multiple links to survive every possible failure.

Paper [25] investigates a very interesting case that, for making the network disaster-resilient, sometimes adding two cables between the same two nodes is cheaper than adding a single cable. This may happen if the edge is relatively short compared to the radius r of the disaster. For example, if the length is shorter than 4 times the radius of the circular disk, it is more economical to deploy two shorter cables than a single (but longer) one. It is proved in [25, Theorem 1.] that any optimal solution of the problem of [25] contains either one or two newly added edges. However, adding three can never be the optimal solution.

As it turns out, adding a single cheapest link that provides *maximal protection* in the worst case is solvable in polynomial time, given that the disaster comes in the form of a circular disk of radius r, and the input graph is planar and connected. Here, maximal protection is defined as follows. The epicenter of a disaster whose failure the network will not survive[2] is called a danger point. All danger points on the plane are partitioned into danger zones, and each new link protects some of the danger zones. Paper [25] proposes a heuristic that adds new links until every danger zone is protected. Nevertheless, once we have selected the danger zones to protect, adding a single link can be done in polynomial time. Note that the complexity of finding the optimal solution with two new cables is an open question.

Proposition 1 (Claim 4 of [25]) *We are given a connected Planar Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with links considered as polygonal chains, a uniform cable cost function, a radius r, a single danger zone Z and two nodes s, t of $\mathcal{G}$. Finding a minimum cost edge between the given nodes with maximal protection for danger zone Z can always be done in polynomial time. The new*

[2] According to Definition 2.

link provides maximal protection for a danger zone if it cannot be hit by a circular disk failure with an epicenter in the danger zone Z and radius r. Note that the distance of both nodes s, t to Z is at least r; otherwise, the problem has no solution.

The algorithm is based on computational geometry algorithms, such as offsetting areas bounded by line segments and circular arcs and the geometric Dijkstra algorithm. Note that the geometric Dijkstra algorithm builds on the assumption that the cost of implementing a new cable is uniform on the plane.

There are algorithms implemented on GPU using OpenGL shaders [5], that can efficiently handle the more general case where this cost varies over the plane. Furthermore, the areas that need to be offset are bounded by line segments and circular arcs. This is because, in the worst-case scenario, the failed areas are assumed to be circular disks. The algorithm needs to offset any arbitrary area in a more general failure model, where a disaster may damage any arbitrary area.

3.2 | Multiple New Links in the Worst-Case Models

If the graph is planar and connected, we have point failures, and multiple new edges are allowed, we have good news:

Proposition 2 (Theorem 3 of [2]) *For a connected Planar Straight Line Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and a weight function $w : \binom{\mathcal{V}}{2} \rightarrow \mathbb{R}_{\geq 0}$, an edge set $\mathcal{E}^+$ of minimum weight such that $\mathcal{G}^+(\mathcal{V}, \mathcal{E} \cup \mathcal{E}^+)$ is 2-connected can be computed in $O(|\mathcal{V}|^4)$ time.*

Note that the 2-connectivity of the constructed graph ensures that the network survives any single point failure (remains connected). We also have a tight bound on the maximum total length of new cables:

Proposition 3 (Theorem 2 of [2]) *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected Planar Straight Line Graph with $|\mathcal{V}| \geq 3$ and no three collinear vertices. Then G can be augmented with straight line segment edges such that the resulting graph $\mathcal{G} = (\mathcal{V}, \mathcal{E} \cup \mathcal{E}^+)$ is planar and 2-connected with $c(\mathcal{E}^+) < 2c(\mathcal{E})$, and this bound is the best possible.*

For disconnected graphs and point failures, the problem becomes $\mathcal{NP}$-hard:

Proposition 4 (Theorem 5 of [2]) *Given is a (disconnected) Planar Straight Line Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and a positive integer w. It is $\mathcal{NP}$-complete to decide whether there exists a set of line segment edges $\mathcal{E}^+$ such that $c(\mathcal{E}^+) \leq w$ and $\mathcal{G}^+ = (\mathcal{V}, \mathcal{E} \cup \mathcal{E}^+)$ is planar and 2-connected.*

Unfortunately, it is not clear how to generalize the proof of the last theorem to regional failures, and to edges that are not necessarily straight line segments between their endpoints. Thus, in the following, we build our proofs of $\mathcal{NP}$-hardness for problems NA, GNA, NE, and GNE based on a reduction to a different $\mathcal{NP}$-hard problem. Namely, we will rely on the $\mathcal{NP}$-completeness of finding a Hamiltonian circuit in *grid graphs* (see also Figure 1):

Definition 3 *$\mathcal{G}(\mathcal{V}, \mathcal{E})$ is a grid graph if its nodes are embedded in the plane with integer coordinates in a related Cartesian coordinate system, and there is an edge between two nodes u and v exactly if their distance $d(u, v) = 1$.*

Claim 1 (Theorem 2.1. of [12]) *The Hamiltonian circuit problem for grid graphs is $\mathcal{NP}$-complete.*

We call the above problem HCGG. See Figure 1 for an example grid graph with a Hamiltonian circuit.

Lemma 5 *The Network Augmentation (NA) problem is $\mathcal{NP}$-hard.*

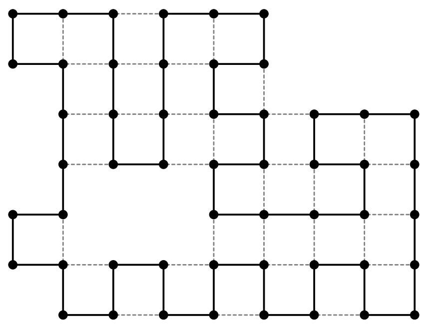


FIGURE 1 An example of a grid graph that contains a Hamiltonian circuit. Edges used by the cycle are drawn with thick solid lines, while the rest of the links are dashed.

Proof We start with a set of nodes $\mathcal{V}$ that have integer coordinates according to a Cartesian coordinate system over the Euclidean plane, and let $|\mathcal{V}| \geq 3$. Let $\mathcal{E}$ be empty, that is, we have not yet built network links. The maximum allowed length of newly installed cables B is set to $n = |\mathcal{V}|$.

We first observe that, to survive a single point failure, the network has to remain connected after the removal of any link or node, i.e., it has to be 2-connected. Thus, the degree of every node in $\mathcal{V}$ has to be at least 2. Since the minimum distance between nodes is 1, this means the cheapest solution to the NA problem has a cost of at least n. Since our budget B is just n, this means that if the NA problem has a solution, it has a cost of B. In addition, the degree of all nodes has to be 2, i.e., the resulting graph has to be a cycle, since we want to build a connected graph.

We claim that any cycle over $\mathcal{V}$ that is not a subgraph of the grid graph induced by $\mathcal{V}$ has a length at least $n+\sqrt{2}-1$, for the following reason. It necessarily contains at least one edge between two nodes that are not adjacent in the grid, and thus are at least $\sqrt{2}$ away from each other. Thus, if there exists a valid solution to our NA problem setting, it can be nothing else than a Hamiltonian circuit $H(\mathcal{V}, E_H)$ in the grid graph induced by $\mathcal{V}$. We note that H does not cross itself either in a geometric sense. Thus, it is a valid solution to the NA problem.

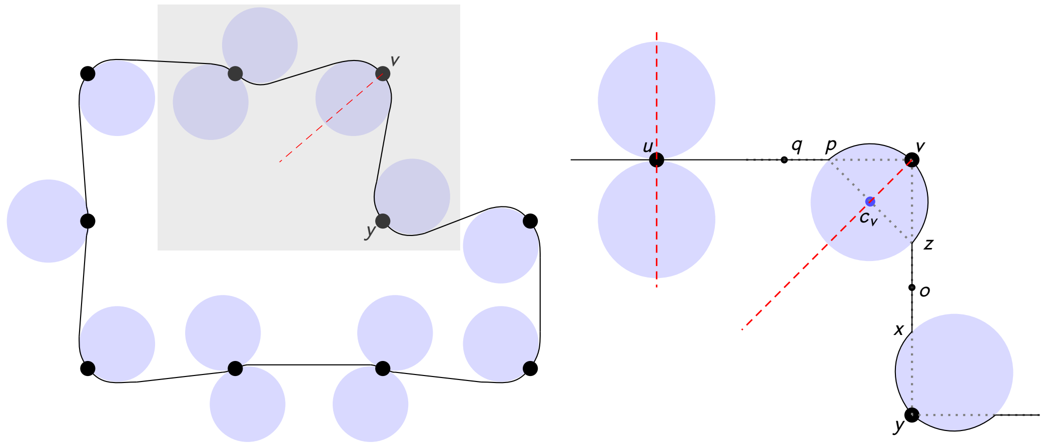
We conclude that the NA problem is $\mathcal{NP}$-hard since the Hamiltonian circuit problem in grid graphs (HCGG) is $\mathcal{NP}$-complete (Claim 1). □

For disconnected graphs and (connected) regional failures the problem remains $\mathcal{NP}$-hard. Here, our most important proposition is that the worst-case problem defined in [25] is $\mathcal{NP}$-hard:

Theorem 6 *The GNA Problem is $\mathcal{NP}$-hard.*

Proof As defined, in the case of the GNA problem, we want the network to survive circular disk failures of a given radius. Compared to this, intuitively, it is easier to cope with single-point failures. Thus, we will blow up the point failures to circular disk failures, and augment the proof of Lemma 5 to cope with this problem too.

Let our GNA problem instance be the following. We only consider the problem on grid graphs, i.e., let the coordinates of every node $v \in \mathcal{V}$ be integers in a related Cartesian coordinate system in the plane (cf. Definition 3) and let $|\mathcal{V}| \geq 3$. Let the initial set of edges $\mathcal{E}$ be empty, that is, we have not yet built network links. Let the disaster radius be $0 \leq r < \frac{1}{n}\frac{\sqrt{2}-1}{\pi-2\sqrt{2}}$. Finally, let the cost budget be $B = \left(n+\sqrt{2}-1+n+nr(\pi-2\sqrt{2})\right)/2$. Note that $B \in \left[n+nr\left(\pi-2\sqrt{2}\right),\ n+\sqrt{2}-1\right)$.



(a) An example grid network with (near) optimal cable routes.

(b) Part of Figure 2 inside the gray rectangle but using the estimates described in the proof of Theorem 6.

FIGURE 2 An example grid network with (near) optimal cable routes as an illustration of the network used in the proof of Theorem 6. Our simplified solution for the overestimation of the optimal cable length in the gray area is depicted in Figure 2b.

Our first observation is that any solution S of the GNA problem is a solution of the NA problem too, since S has to remain connected in case of any point failure. This observation, combined with the proof of Lemma 5 yields that any solution $S(\mathcal{V}, E_S)$ of our GNA problem instance (that has to have a cost at most $B < n + \sqrt{2} - 1$) has to have the following properties:

1. S is a cycle,
2. every edge in E_S connects two nodes that are adjacent in the grid.

This means that if there exists a solution to our GNA problem setting, then (neglecting the geometric embedding for a moment) it has to be a Hamiltonian circuit over the grid graph induced by $\mathcal{V}$. To complete our argument of $\mathcal{NP}$-hardness of the GNA problem, we need to show the following: if there exists a Hamiltonian circuit on the grid graph induced by $\mathcal{V}$, then there exists a solution of our problem setting with cost $\leq B$, where S is called Hamiltonian-circuit-based graph.

In other words, we will show that we can use curves for the edges of the Hamiltonian circuit such that the network survives any disk failure. Moreover, for small enough r, this solution has a length less than $n + \sqrt{2} - 1$.

The high-level idea is that, in the case of an optimal solution, in the vicinity of each node, the curve is a circular arc of radius r. These arcs are connected by straight line segments, *without any corner points* (see Figure 2). One can imagine it as if there are circular disks (e.g., coins) attached to each side of the cable at each node. We pull the cable to be as short as possible (as if it was a rubber band) since the goal is to minimize the total cable length. In this case, the cable will traverse each node along a circular arc; otherwise, it consists of line segments.

Unfortunately, the exact cable routes in an optimal solution are extremely difficult to characterize, since the local structure of the curve is heavily dependent on the exact layout of the rest of the topology. Thus, instead of the shortest solution, we will study a class of slightly suboptimal solutions. In these solutions, the route (and thus the length) of the cable in the neighborhood of each node is independent of the route of the cable around other nodes.

Instead, it depends solely on the disaster radius.

Our main idea is that for a GNA problem instance on a node set $\mathcal{V}$ that has a Hamiltonian circuit in its induced grid graph, we modify the initial straight line segment graph realization only at the close vicinity of the nodes. This approach enables us to study the total edge length of the resulting graph much more easily.

More concretely, in our approach, we restrict the possible places of the center points c_w of the imaginary coins attached to each node w (depicted in Figure 2, e.g. node v and the corresponding center point c_v). The center points can be on the grid lines (at nodes where the Hamiltonian circuit follows a straight path, like node u on Figure 2b). Alternatively, they can be positioned at a diagonal of a cell (e.g., straight line vc_v is in a 45° direction compared to the vertical line on Figure 2b). We call nodes where the Hamiltonian circuit follows a straight path as *non-turning node*s, while the rest as *turning node*s. For a turning node w, we call the diagonal straight line wc_w described above as the *diagonal of (turning node) w*.

For simplicity in our calculations, at each turning node w of the circuit we require the link routes to follow the perimeter of the coin (circular disk centered at c_w, and having radius r) until it reaches the next intersection point with the grid (like points p and z for node v on Figure 2b). Then we require them to cross along the grid lines until they bump into another coin. This way, at each turning node, we force the links to travel around an arc of 180°, which translates to a cable length of πr. With straight line segments ignoring the coins, this route would have been $2r\sqrt{2}$.

We note that this setting yields a valid solution to the GNA problem, because of the following. For a non-turning node w', it is trivial that a disaster with radius r either hits node w', or hits at most one of its two incident edges. For a turning node w, the center point c of a circular-disk-shaped disaster D of radius r is either 1) on the diagonal of w, or 2) on either side of the diagonal. If c is on the diagonal, then it hits both links incident to w exactly if w is hit too. If c is not on the diagonal of w, then it is on either side of the diagonal. It can hit the link incident to w on the other side of the diagonal only if it hits node w too. Thus, there is no circular disk of radius r whose failure separates the network into multiple components.

We can see that each turning node in the solution adds $(\pi - 2\sqrt{2})r$ to the original total cable length of n of the Hamiltonian circuit. Since there are at most n turning nodes, this solution has a total length $c(\mathcal{E}^+) \leq n\left(1 + (\pi - 2\sqrt{2})r\right)$. Thus, for disaster radii $r < \frac{1}{n}\frac{\sqrt{2}-1}{\pi-2\sqrt{2}}$, the total cable length $c(\mathcal{E}^+) < B$. Based on the above, we conclude that the GNA problem is $\mathcal{NP}$-hard, since the Hamiltonian circuit problem in grid graphs is $\mathcal{NP}$-complete (by Claim 1). □

3.3 | Multiple New Links and Nodes in the Worst-Case Models

Here we only have $\mathcal{NP}$-hardness results, which are only for disconnected graphs.

The above problems assumed that only new cables could be installed. However, installing new network nodes along new cables is also an option. With this in mind, we now turn to present the proofs of $\mathcal{NP}$-hardness of the NE and GNE problems. First, we present the following Claim 2 that is a rephrasing of [17, Lemma 2]. The original lemma is tackling the *2-Connected Steiner Network Problem in the Plane (2SNPP)* [17], that, in the case of no network edges (i.e., $\mathcal{E}$ is empty), is a special case of our Network Extension problem.

Claim 2 (rephrased Lemma 2 of [17]) *Let $\mathcal{V}$ be a set of n integer grid points, and let the set $\mathcal{E}$ of original links be empty. Then, we have the following.*

- *Every solution to the NE problem is of length at least n.*
- *The only solutions to the NE problem of length exactly n are Hamiltonian circuits such that every two successive points are adjacent in the grid.*

- *All other solutions are of length at least* $n + \sqrt{2} - 1$.

Lemma 7 *The Network Extension (NE) problem is* $\mathcal{NP}$*-hard.*

Proof We start with a set of nodes $\mathcal{V}$ with integer coordinates according to a Cartesian coordinate system over the Euclidean plane, and let $|\mathcal{V}| \geq 3$. Let $\mathcal{E}$ be empty, i.e., we have no network links built yet. The maximal installation budget B is set to n. Since the cost of installing new nodes $c(\mathcal{V}^+) \geq 0$, we have to satisfy $c(\mathcal{E}^+) \leq n$ (the cost of installing new edges cannot exceed n). By Claim 2, if there exists a solution to this Network Extension problem setting, it has to be a Hamiltonian circuit that is a subgraph of the grid graph induced by $\mathcal{V}$. We conclude that the NE problem is $\mathcal{NP}$-hard, since the HCGG problem is $\mathcal{NP}$-complete (Claim 1). □

Lemma 8 *The Geometric Network Extension (GNE) problem is* $\mathcal{NP}$*-hard.*

Proof Let the settings of our GNE problem instance be similar to those seen in the proof of Theorem 6. More precisely, we consider the following. We only consider the problem on grid graphs, i.e., let the coordinates of every node $v \in \mathcal{V}$ be integers in a related Cartesian coordinate system in the plane, and let $|\mathcal{V}| \geq 3$. Let $\mathcal{E}$ be empty, i.e., we have no network links built yet. Let the disaster radius be $0 \leq r < \frac{1}{n}\frac{\sqrt{2}-1}{\pi-2\sqrt{2}}$. Finally, let maximal network extension budget be $B = \left(n + \sqrt{2} - 1 + n + nr(\pi - 2\sqrt{2})\right)/2$. Note that $B \in \left[n + nr\left(\pi - 2\sqrt{2}\right), n + \sqrt{2} - 1\right)$.

We observe that a solution to the GNE problem is a solution to the NE problem too. This, combined with Claim 2 means that a valid solution of the GNE problem with a cost smaller than $n + \sqrt{2} - 1$ has to visit the vertices of $\mathcal{V}$ according to a Hamiltonian circuit on the grid graph induced by $\mathcal{V}$. Note that if a Hamiltonian-circuit-based solution described above exists, it has a cost of $n + nr(\pi - 2\sqrt{2}) < B$. Also, since the cost of adding new nodes $c(\mathcal{V}^+) \geq 0$, there would be no benefit in adding new nodes.

We conclude that the GNE problem is $\mathcal{NP}$-hard, since the Hamiltonian circuit problem in grid graphs is $\mathcal{NP}$-complete (Claim 1). □

3.4 | Further Observations on the Complexity of Disaster-Aware Network Augmentation and Extension

In our proofs of $\mathcal{NP}$-hardness of GNA and GNE problems, we could state that the cost of an optimal Hamiltonian-circuit-based solution is strictly less than $n + \sqrt{2} - 1$ if the disaster radius $r < \frac{1}{n}\frac{\sqrt{2}-1}{\pi-2\sqrt{2}}$. Since $1.32 < \frac{\sqrt{2}-1}{\pi-2\sqrt{2}}$, we can state the following[3].

Corollary 9 *The Geometric Network Augmentation (GNA) and Geometric Network Extension (GNE) problems are* $\mathcal{NP}$*-hard on the node sets* $\mathcal{V}$ *of grid graphs, and a disaster radius* $r \leq 1.32/|\mathcal{V}|$.

Proof The claim immediately follows from the proofs of Theorem 6 and Lemma 8. □

As we have seen, all the problems NA, GNA, NE, and GNE are $\mathcal{NP}$-hard. Looking inside their proofs of $\mathcal{NP}$-hardness, we can see that there is even no Fully Polynomial Time Approximation Scheme (FPTAS) for these problems.

Corollary 10 *For each of the NA, GNA, NE, and GNE problems, there is no Fully Polynomial Time Approximation Scheme (FPTAS), unless* $\mathcal{P} = \mathcal{NP}$*. More precisely, no polynomial-time* $1+1/5n$ *approximation exists for these problems, unless* $\mathcal{P} = \mathcal{NP}$.

[3] Note that the maximal radii for which our arguments on $\mathcal{NP}$-hardness hold are bigger than $\simeq 1.32/n$, by a (small) constant factor. However, since our calculations for optimizing this constant were overly complicated, we decided to exclude them from this study.

Proof An FPTAS takes an instance of a problem (e.g., NA, GNA, NE, or GNE) and a parameter $\epsilon > 0$ as input. In the case of minimization problems, it returns as output a solution whose value is at most $1 + \epsilon$ times the optimum within a run-time that is polynomial in the problem size and in $1/\epsilon$. A consequence of the following argument is that in the case of NA, GNA, NE, or GNE, no such algorithm exists for $\epsilon = 1/5n$ unless $\mathcal{P} = \mathcal{NP}$.

For an instance of the NA and NE problems, we have from Lemma 5 and 7, respectively, that a solution to the problem instance that is a Hamiltonian circuit on the grid graph induced by $\mathcal{V}$ has a cost of n. Any other solution must have length at least $n + \sqrt{2} - 1 > n + 0.414$. Now for $\epsilon = 0.414/n$, an ϵ-approximation to the NA and NE problems, respectively, will have a length less than $n + 0.414$ if and only if the HCGG problem has a feasible solution. It follows that there can be no algorithm that finds an ϵ-approximation in time polynomial in $1/\epsilon = n/0.414$ unless $\mathcal{P} = \mathcal{NP}$.

Similarly, in the case of geometric problems, Hamiltonian-circuit-based solutions have a cost of at most $n + nr(\pi - 2\sqrt{2})$. For $r = \frac{1}{n}\frac{\sqrt{2}-1}{\pi-2\sqrt{2}}$ this length is $n + \frac{\sqrt{2}-1}{2}$. All the other solutions cost at least $n + \sqrt{2} - 1$. Now for $\epsilon = \frac{\sqrt{2}-1}{2}/n \simeq 0.207/n$, an ϵ-approximation to the GNA and GNE problems, respectively, will have a length less than $n + 0.207$ if and only if there exists a Hamiltonian circuit in the grid graph induced by $\mathcal{V}$. It follows that there can be no algorithm that finds an ϵ-approximation in time polynomial in $1/\epsilon \simeq n/0.207$ unless $\mathcal{P} = \mathcal{NP}$. Finally, $0.207/n > 1/5n$, which completes the proof. □

We note that the inexistence of an FPTAS does not mean that big instances of the problems cannot be solved efficiently and nearly optimally. In fact, indifferently of n, the best Hamiltonian-circuit-based solutions are no more than $\simeq 0.414$ shorter than any other type of solution. The difference represents only an additive gap, which may be negligible in the case of large networks extending over vast areas.

3.5 | On the Hardness of the Horizontal Line Segment Failure Model

In the current subsection, we give our thoughts on the problem discussed in [3]. Informally speaking, Claim 3 proves the $\mathcal{NP}$-hardness of the GNA and GNE problems in the case of horizontal line segment failures with length $r < 1$.

Claim 3 *Given a network represented by an undirected geometric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Nodes are embedded as points, and edges as line segments in the Euclidean plane. Given the maximum length r of a possible horizontal line segment failure, and the limit B on the total cost budget of installing new network equipment. B is the sum of the total length of curves $c(\mathcal{E}^+)$ and the installation cost of new nodes $c(\mathcal{V}^+)$. Then, it is $\mathcal{NP}$-hard to decide whether there is a node set $\mathcal{V}^+$ added to $\mathcal{V}$ and curves $\mathcal{E}^+$ through $\mathcal{V} \cup \mathcal{V}^+$ added as new edges to $\mathcal{E}$, such that the network survives any horizontal line segment failure of length r, and the total cost of network extension is at most B. The problem remains $\mathcal{NP}$-hard if we forbid new nodes.*

Proof We only consider the problem on grid graphs. Let $\mathcal{E}$ be empty. Let the length of the disaster line segment be $r < 1$. Finally, let the budget be $B = \left(n + \sqrt{2}-1/2\right)$.

We claim that both problems stated in the claim (GNA and GNE) have solutions exactly if the NA problem has a solution on the same node set $\mathcal{V}$ and budget $B = \left(n + \sqrt{2}-1/2\right)$. If the NA problem has a solution, we take the one that uses straight line segments (i.e., the grid line segments) to connect the nodes. This claim holds simply because any horizontal line segment failure of length less than 1 hits at most 1 node of $\mathcal{V}$, and the edges hit by this line segment are incident to the same node. The proof follows. □

We note that another way of proving this claim for sufficiently short horizontal line segments ($r < \frac{2}{n}\frac{\sqrt{2}-1}{\pi-2\sqrt{2}}$) can be done via overestimating the line segment with a circular disk with diameter r, and simply referring to Theorem 6 and Lemma 8.

The problem formulation of Claim 3 where no new nodes are allowed is the same as the Line Segment Disaster Augmentation Problem (LSDAP) presented in [3], but without requiring $\mathcal{G}$ to be connected. As a corollary, we can state that a version of the LSDAP where the input graph does not have to be connected is $\mathcal{NP}$-hard. This is a stronger indication towards possible $\mathcal{NP}$-hardness of the (general) LSDAP than Proposition 4 (Theorem 5 of [2]), since here, the length of the line segment failures are allowed to be strictly positive.

At the same time, we can see that in the case of the grid graphs and horizontal line segment failures shorter than the unit, the LSDAP is not harder than any of the problems NA, GNA, NE, and GNE. Finally, the authors of [3] indicate that the LSDAP has similarities with the 2-connected augmentation of planar graphs. The latter is solvable in polynomial time if both existing and new network links are line segments. In conclusion, despite our complexity results based on the $\mathcal{NP}$-hardness of finding a Hamiltonian circuit in a grid graph, proving the conjectured $\mathcal{NP}$-hardness of the LSDAP (on connected graphs) is still a challenging task.

4 | PROBLEMS BASED ON COST-EFFICIENT MODELS

Cost-efficient models assume that the probabilities of the disasters are also known. They take a more comprehensive goal than searching among the worst-case scenarios. The aim is to find a **cost-efficient** investment cost as a trade-off with the risk of a network outage [22]. The formulated problems are $\mathcal{NP}$-hard, because a lot of historical hazard data must be processed, and heavy algorithms must be used to find a cost-efficient solution. In this section, we briefly overview the main characteristics of selected worst-case models.

In [18], a disaster-aware submarine cable deployment algorithm was devised by exploring a greenfield network planning approach. The primary objective of this approach is to prevent the network infrastructure from being deployed in the disaster areas. The study assumes a set of candidate new cables are given and provides an Integer Linear Program (ILP) to select the set of new links with minimal cost.

In [31], a cost-effective approach was presented for planning submarine cable routes (paths) with minimized overall life-cycle cost of the submarine cables. By assuming irregular shapes of the cable routes, an Integer Linear Program (ILP) was formulated for physical path selection. The optimality of the ILP is nonetheless affected by some additional constraints. Note that the study did not consider planning several paths simultaneously.

The study [22] assumes that a disaster destroys some points in the plane. If a link or node intersects with any of these disaster points, it is destroyed. The set of disasters is predefined ($\sim 500\,000$ in the simulations), and each disaster can destroy an arbitrary set of points in the plane. Note that the model is very general, as **it does not even assume that the points of a disaster must be connected**. For each predefined disaster, a probability is also assigned, such that these probabilities are summed up to 1 in the input. This failure model is similar to Probabilistic Shared Risk Link Group failures, where the set of failure events with their probabilities is known in network planning [27, 29, 21, 1, 15].

In the disaster model of [22], it is challenging to define whether a new cable is affected by a disaster if it can take any arbitrary route. To mitigate this problem, [22] limits the route of the cable to traverse along a 2D grid. More precisely, the cable should traverse the points of a predefined grid, where it can take a vertical, horizontal, or diagonal line segment, see the dashed (blue) link on Figure 3 as an example. The cost of implementing a cable strongly depends on the location, which can be easily included in the grid model. On the other hand, this restriction comes with certain disadvantages. The problem input size strongly depends on the size of the cells. The authors evaluate 5.5km cells (0.05 degree) over Italy in the simulations. Another drawback of mapping the routes to the grid is that it introduces an error in its length evaluation. Compared to Euclidean distances, the excess length can be as much as 8%, see Section 5.

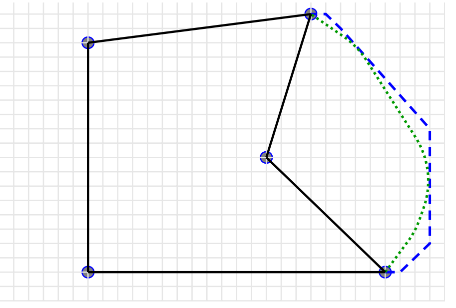


FIGURE 3 The dashed (blue) line segments illustrate the construction of a new edge in the cost-efficient model of [22]. The edge is along a predefined grid, where it can take a vertical, horizontal, or diagonal line segment. The dotted (green) curve demonstrates the shape of the added cables in a worst-case model (like the model of [25]). There, it avoids a circular area to survive disk failures.

Note that the failure probabilities of [22] are computed using stochastic models of the disasters, where the failed points form circular disks in the simulations of [27, 29]. On the contrary, the network should be prepared for an infinite number of possible disaster scenarios in the worst-case analysis. In other words, the probability of each disaster is ignored, and the solution should survive any possible disaster. This is a fundamentally different concept that can be solved without restricting the routes of the cables to a grid; see the dotted (green) curve in Figure 3 as an example.

The cost-efficient model in [22] is more general than the worst-case models in two aspects.

1. It defines an impact metric (denoted by M and similar to a traffic matrix). M defines how important it is to avoid losing a connection between a given source and destination node pair in case of disasters.
2. It defines the cost of losing a connection between two nodes because of a disaster (denoted by α) relative to the cost of implementing 1 km optical fiber.

The objective function is composed of two parts: the first is α times the impact of losing a connection because of a disaster, while the second is the cost of implementing new cables. In the simulation of [22], M is the number of disconnected node pairs divided by the total amount of node pairs. Thus, $M = 0$ if all node pairs are connected, and $M = 1$ if all node pairs are disconnected. Note that the worst-case model is equivalent to $\alpha = \infty$. In the worst-case model, $M = 0$ if all node pairs are connected, and $M > 0$ otherwise. Note that the value of M is not relevant in worst-case models like [25], since $\alpha = \infty$.

The paper [22] deals with complexity results as well. It turns out that adding a single cable in their cost-efficient model is $\mathcal{NP}$-hard (Theorem 1 of [22]). The construction in the $\mathcal{NP}$-hardness proof in Theorem 1 of [22] builds on the existence of disasters that result in a simultaneous failure of multiple network elements very far from each other. It is not known whether the problem remains $\mathcal{NP}$-hard if every disaster has to be a connected set of points.

5 | LIMITATIONS OF GRID-BASED CABLE ROUTES

In this section, we evaluate the maximum error in the cost function because of limiting the routes of the new cables to connect a series of neighboring cell center points. We call this effect the discretization of the topology map.

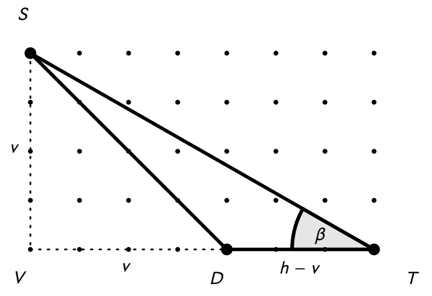


FIGURE 4 Illustration for the proof of Lemma 11, where the points denote the center points of the grid's cells. $\overline{ST}$ denotes the Euclidean shortest path between S and T. $\overline{SD} + \overline{DT}$ denote the shortest grid-based cable route. The grid-based cable route can be up to $\sim 8.24\%$ longer than the Euclidean distance. The greatest error occurs at $\beta = \pi/8$.

Lemma 11 *Suppose we have a grid where the cells have unit side lengths, and each endpoint of the cable is placed in the center of a cell. Points can be connected using vertical, horizontal, or diagonal line segments that connect adjacent cells in the grid. Here, two cells are said to be adjacent if they share a common face or corner; consequently, each cell has* 8 *neighbors. Suppose we want to connect two cell center points S and T via a chain of neighboring cell center points. Then, the resulting route can be* $\sim 8.24\%$ *more expensive than line segment $\overline{ST}$, for uniform cable cost.*

Proof Let v be the vertical, and h be the horizontal distance between endpoints S and T. Generally, if we place a horizontal and a vertical line segment after each other, the same endpoint can be reached by a shorter diagonal line segment. Thus, as the goal is to minimize costs, cables cannot contain both horizontal and vertical line segments. We will use this fact to construct the shortest path.

Starting from S, first, we take v vertical, then h horizontal steps towards T. This path connects S and T, but not in the shortest way possible. We take one horizontal and one vertical line segment, and substitute these two with one diagonal line segment. Then we repeat this step as many times as possible. Afterwards, we will have $\min(v,h)$ diagonal, and $|h - v|$ either horizontal or vertical line segments: horizontal if $v < h$, vertical if $v > h$, and neither if $v = h$.

Starting from S, we place these line segments after each other, such that each added line segment brings the cable's endpoint closer to T (otherwise, the path would not be the shortest possible). This way, the last line segment successfully reaches T. Note that the order of the line segments is irrelevant, as we reach the same endpoint T. We consider the construction where the diagonal line segments are placed first, forming one longer line segment. The vertical/horizontal line segments are added, again forming one longer line segment. For example, see $\overline{SD} + \overline{DT}$ in Figure 4. Each cable that contains a vertical line segment can be converted to a cable with a horizontal line segment by rotating it about point S, by $\pi/2$. After the rotation, the length of the cable and the Euclidean distance between S and T remain the same. Therefore, it is enough to consider solutions without a vertical line segment (just as in Figure 4).

To find the greatest possible error, we need to maximize the ratio R between the grid-based and Euclidean cable lengths.

$$R = \frac{\sqrt{2} \cdot v + (h - v)}{\sqrt{v^2 + h^2}} \tag{1}$$

Let β denote the angle between $\overline{ST}$ and the horizontal distance between S and T ($\overline{VT}$). As $v = \tan(\beta) \cdot h$, we can substitute v with $\tan(\beta) \cdot h$ in the above expression.

$$R = \frac{((\sqrt{2} - 1)\tan\beta + 1) \cdot h}{\sqrt{(\tan^2\beta + 1) \cdot h^2}} \tag{2}$$

As $h > 0$ and $\cos\beta > 0$ for all possible angles in a triangle, $\sqrt{\tan^2\beta + 1} = \frac{1}{\cos\beta}$, and the value of R is the following.

$$R = (\sqrt{2} - 1)\sin\beta + \cos\beta \tag{3}$$

To find R, we need to find β such that $\dot{R}$, the value of the first derivative with respect to β is 0.

$$\dot{R} = (\sqrt{2} - 1)\cos\beta - \sin\beta = 0 \tag{4}$$

Knowing that $0 < \beta < \pi/2$, the only solution in this domain is exactly at $\beta = \pi/8$.

From triangle SVD we know that angle $\measuredangle(VDS) = \pi/4$, therefore angle $\measuredangle(SDT) = 3\cdot\pi/4$. As $\beta = \pi/8$, angle $\measuredangle(DST)$ also has to be $\pi/8$. Therefore SDT is an isosceles triangle, and $h = (\sqrt{2} + 1) \cdot v$. Thus, if $\beta = \pi/8$,

$$R = (\sqrt{2} - 1)\sin\pi/8 + \cos\pi/8 \approx 1.0824. \tag{5}$$

Note that v and h are integers because they connect the center points of grid cells. There are no possible integer values for $h = (\sqrt{2} + 1) \cdot v$, therefore, in practice, the exact maximum (more precisely, supremum) error cannot be reached. On the other hand, for instance, if $h = 169$ and $v = 70$, conversion to the grid results in a close approximation to the maximum error, with only a difference of $\sim 6 \cdot 10^{-11}$. □

As indicated in [22], we note that this maximum error can be decreased by enabling additional cable directions. However, it further complicates the optimization problem.

6 | CONCLUSION

This paper focuses on how to extend a network topology to make it more resilient against disasters. We investigated different mathematical models and formulated problems, each capturing a unique aspect. First, we overviewed the problems that can be solved with polynomial algorithms. Next, we showed that four Network Extension/Augmentation problems are $\mathcal{NP}$-hard. They require adding new cables to the network until 100% survival rate is achieved against natural disasters in the worst case. We also showed that these problems remain $\mathcal{NP}$-hard if nodes can be added besides the new links. Furthermore, we discussed the limitations of a more sophisticated problem, where the goal is to find the minimum-cost extension as a compromise with the level of protection against disasters. In this case, detailed input data, such as historical data on the disasters and the cost of implementing new cable routes at each location, is needed. The model requires the new cable routes to follow a grid, because of the discretization of the topology map. We showed that this restriction results in a possibly significant suboptimality of the result, no matter how precise the input data is. All in all, the paper is a further step in understanding what network-extending problems can be solved efficiently.

References

[1] P.K. Agarwal, A. Efrat, S.K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, *The resilience of WDM networks to probabilistic geographical failures*, IEEE/ACM Trans. Net **21** (2013), 1525–1538.

[2] H.A. Akitaya, R. Inkulu, T.L. Nichols, D.L. Souvaine, C.D. Tóth, and C.R. Winston, *Minimum weight connectivity augmentation for planar straight-line graphs*, Theor. Comput. Sci. **789** (2019), 50–63, Selected Papers from the 11th International Conference and Workshops on Algorithms and Computation.

[3] N. Andrés-Thió, M. Brazil, C. Ras, and D. Thomas, *Network augmentation for disaster-resilience against geographically correlated failure*, Networks (2022).

[4] F. Dikbiyik, M. Tornatore, and B. Mukherjee, *Minimizing the risk from disaster failures in optical backbone networks*, J. Lightwave Technol. **32** (2014), 3175–3183.

[5] R. Farias and M. Kallmann, *Optimal path maps on the gpu*, IEEE transactions visualization computer graphics **26** (2019), 2863–2874.

[6] M.R. Garey and D.S. Johnson, *Computers and intractability: A guide to the theory of NP-completeness*, 1978.

[7] O. Gerstel, M. Jinno, A. Lord, and S.B. Yoo, *Elastic optical networking: A new dawn for the optical layer?*, Commun. Magazine, IEEE **50** (2012), s12–s20.

[8] T. Gomes, J. Tapolcai, C. Esposito, D. Hutchison, F. Kuipers, J. Rak, A. de Sousa, A. Iossifides, R. Travanca, J. André, L. Jorge, L. Martins, P.O. Ugalde, A. Pašić, D. Pezaros, S. Jouet, S. Secci, and M. Tornatore, *A survey of strategies for communication networks to protect against large-scale natural disasters*, 8th International Workshop on Resilient Networks Design and Modeling (RNDM), Sept 2016, pp. 11–22.

[9] M.F. Habib, M. Tornatore, M. De Leenheer, F. Dikbiyik, and B. Mukherjee, *Design of disaster-resilient optical datacenter networks*, J. Lightwave Technol. **30** (2012), 2563–2573.

[10] I.B.B. Harter, D. Schupke, M. Hoffmann, G. Carle, et al., *Network virtualization for disaster resilience of cloud services*, IEEE Commun. Ma **52** (2014), 88–95.

[11] J. Heidemann, L. Quan, and Y. Pradkin, *A preliminary analysis of network outages during hurricane Sandy*, University of Southern California, Information Sciences Institute, 2012.

[12] A. Itai, C.H. Papadimitriou, and J.L. Szwarcfiter, *Hamilton paths in grid graphs*, SIAM J. Comput. **11** (1982), 676–686.

[13] S.A. Jyothi, *Solar superstorms: planning for an internet apocalypse*, Proceedings of the 2021 ACM SIGCOMM 2021 Conference, 2021, pp. 692–704.

[14] A. Kwasinski, *Effects of notable natural disasters from 2005 to 2011 on telecommunications infrastructure: Lessons from on-site damage assessments*, Int. Telecommunications Energy Conference (INTELEC), 2011, pp. 1–9.

[15] H.W. Lee, E. Modiano, and K. Lee, *Diverse routing in networks with probabilistic failures*, IEEE/ACM Trans. Networki **18** (2010), 1895–1907.

[16] X. Long, D. Tipper, and T. Gomes, *Measuring the survivability of networks to geographic correlated failures*, Optical Switching Networking **14** (2014), 117–133.

[17] E.L. Luebke and J.S. Provan, *On the structure and complexity of the 2-connected steiner network problem in the plane*, Oper. Res. Lett. **26** (2000), 111–116.

[18] D.L. Msongaleli, F. Dikbiyik, M. Zukerman, and B. Mukherjee, *Disaster-aware submarine fiber-optic cable deployment for mesh networks*, J. Lightwave Technology **34** (2016), 4293–4303.

[19] B. Mukherjee, M. Habib, and F. Dikbiyik, *Network adaptability from disaster disruptions and cascading failures*, IEEE Commun. Ma **52** (2014), 230–238.

[20] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, *Assessing the vulnerability of the fiber infrastructure to disasters*, IEEE/ACM Trans. Net **19** (2011), 1610–1623.

[21] J. Oostenbrink and F. Kuipers, *Computing the impact of disasters on networks*, ACM SIGMETRICS Performance Evaluation Review **45** (2017), 107–110.

[22] J. Oostenbrink and F. Kuipers, *Going the extra mile with disaster-aware network augmentation*, IEEE INFOCOM 2021 - IEEE Conference on Computer Communications, 2021, pp. 1–10.

[23] J. Rak, D. Hutchison, J. Tapolcai, M.T. Rasa Bruzgiene, C.M. Machuca, M. Furdek, and P. Smith, "*Fundamentals of communication networks resilience to disasters and massive disruptions*," *Guide to Disaster-Resilient Communication Networks*, J. Rak and D. Hutchison (eds.), Springer, 2020, pp. .

[24] R. Souza Couto, S. Secci, M. Mitre Campista, K. Costa, and L. Maciel, *Network design requirements for disaster resilience in IaaS clouds*, IEEE Commun. Ma **52** (2014), 52–58.

[25] J. Tapolcai, Z.L. Hajdú, A. Pašić, P.H. Ho, and L. Rónyai, *On network topology augmentation for global connectivity under regional failures*, Proc. IEEE INFOCOM, Vancouver, Canada, May 2021.

[26] P.N. Tran and H. Saito, *Geographical route design of physical networks using earthquake risk information*, IEEE Commun. Magazine **54** (2016), 131–137.

[27] A. Valentini, B. Vass, J. Oostenbrink, L. Csák, F. Kuipers, B. Pace, D. Hay, and J. Tapolcai, *Network resiliency against earthquakes*, 2019 11th International Workshop on Resilient Networks Design and Modeling (RNDM), Oct 2019, pp. 1–7.

[28] B. Vass, B. Brányi, B.E. Nagy, and J. Tapolcai, *On the complexity of disaster-aware network extension problems*, 2022 12th International Workshop on Resilient Networks Design and Modeling (RNDM), 2022, pp. 1–8.

[29] B. Vass, J. Tapolcai, Z. Heszberger, J. Bíró, D. Hay, F.A. Kuipers, J. Oostenbrink, A. Valentini, and L. Rónyai, *Probabilistic Shared Risk Link Groups Modeling Correlated Resource Failures Caused by Disasters*, IEEE J. Selected Areas Commun. **39** (2021), 2672–2687.

[30] V.V. Vazirani, *Approximation algorithms* Vol. 1, Springer, 2001.

[31] Q. Wang, J. Guo, Z. Wang, E. Tahchi, X. Wang, B. Moran, and M. Zukerman, *Cost-effective path planning for submarine cable network extension*, IEEE Access **7** (2019), 61883–61895.